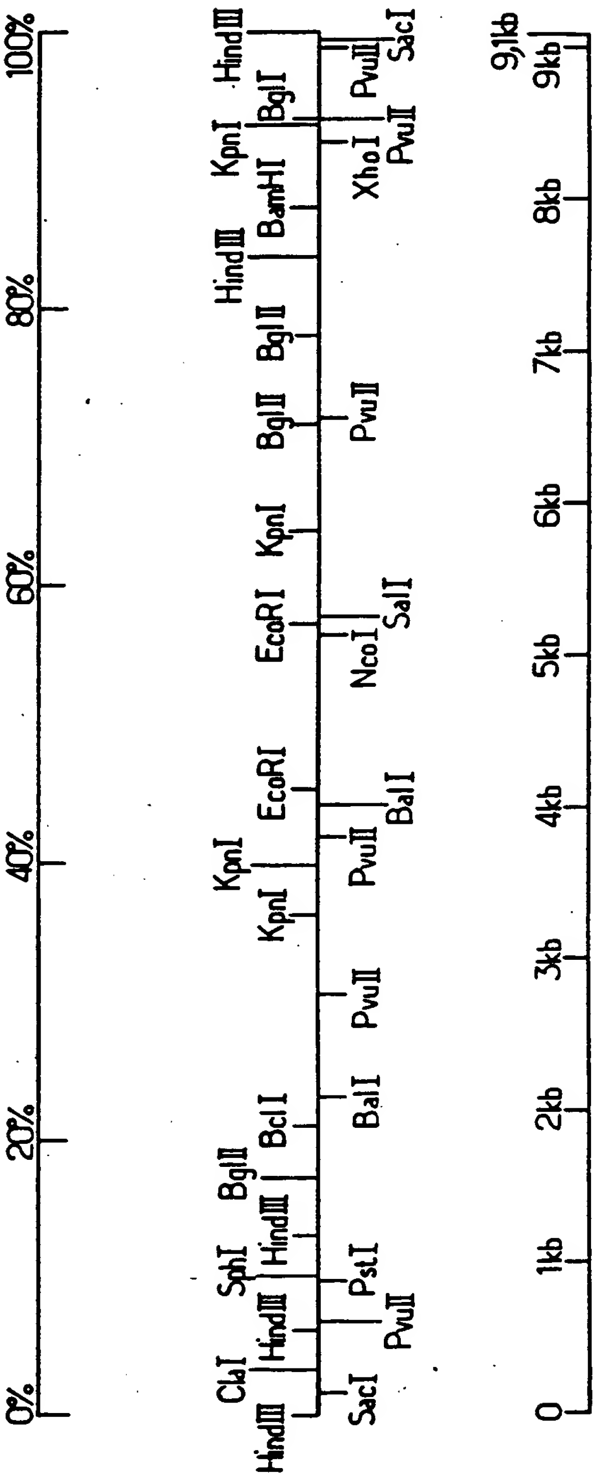
771248

08 026736



0%
20%
40%
60%
80%
100%
HindIII
ClaI
HindIII
SphI
HindIII
BglII
BclI
KpnI
KpnI
EcoRI
EcoRI
KpnI
BglII
BglII
HindIII
BamHI
KpnI
BglI
HindIII
SacI
PvuII
PstI
BalI
PvuII
PvuII
BalI
NcoI
SalI
PvuII
XhoI
PvuII
PvuII
SacI
0
1kb
2kb
3kb
4kb
5kb
6kb
7kb
8kb
9kb
9,1kb

FIG.1.

771248

08 026736

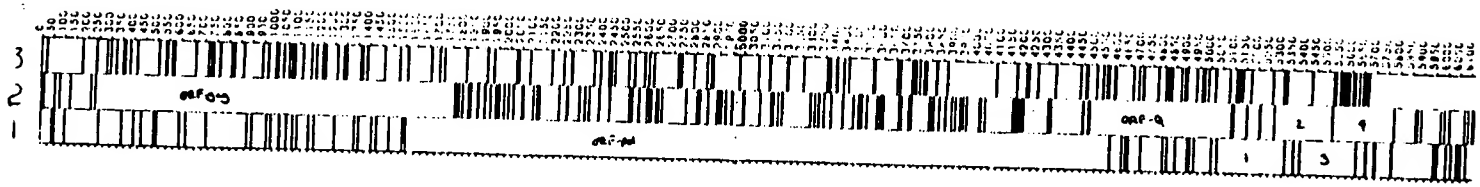

3
2
1
ORF-Q
1
2
3
4

Fig. 2

2d

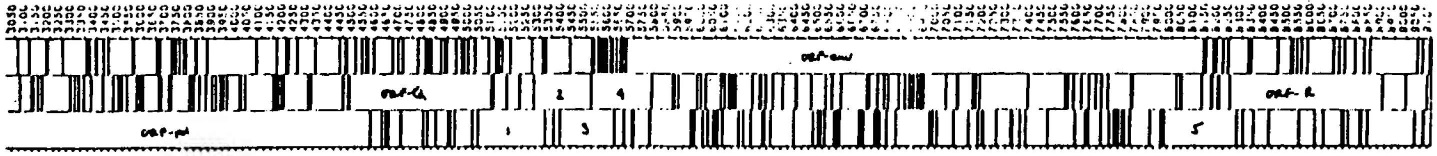


Fig. 3

771248

08 026736

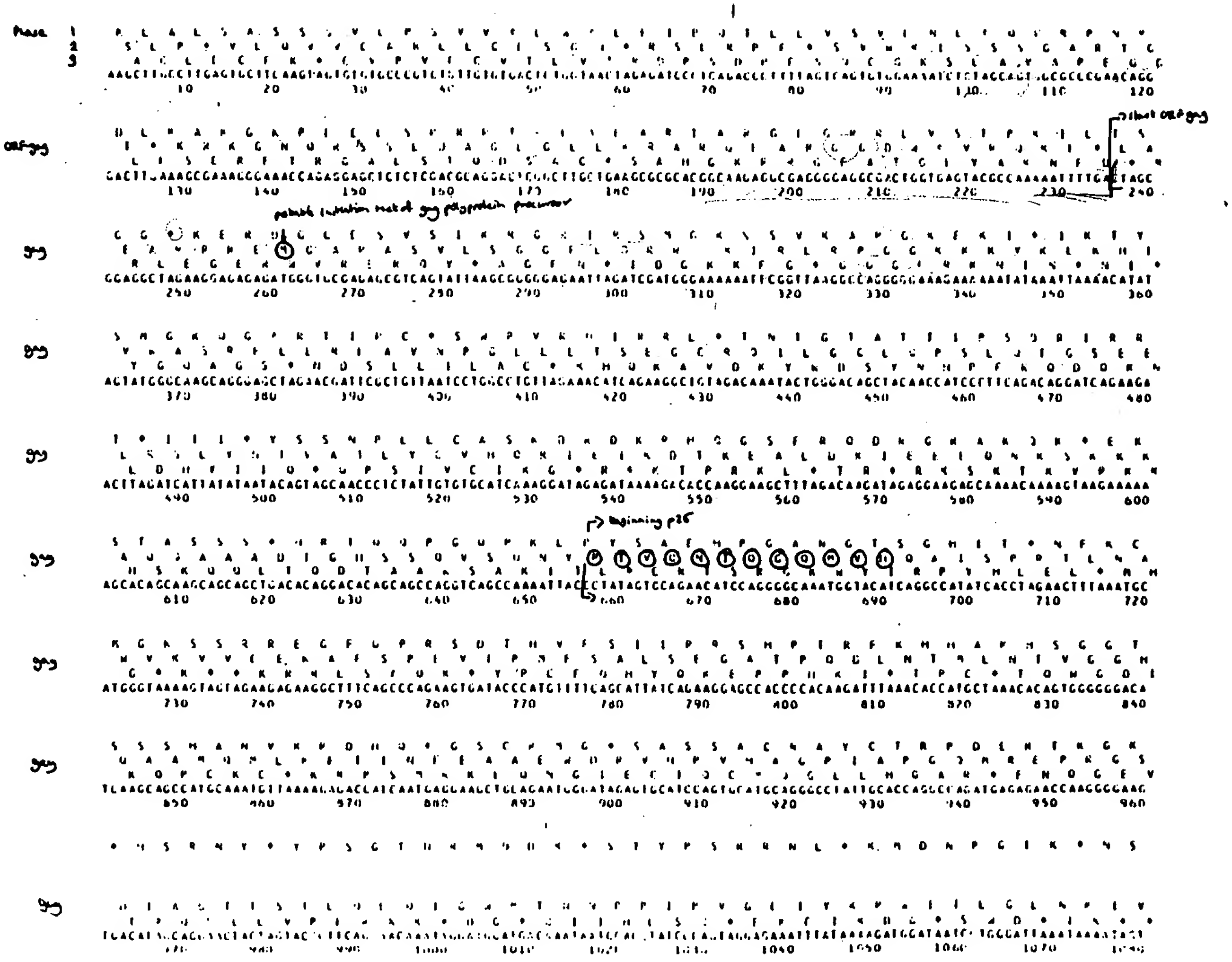

Phase 1
2
3
ORF gag
gag
start ORF gag
potent initiation met of gag polyprotein precursor
beginning p26
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
840
850
860
870
880
890
900
910
920
930
940
950
960
970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080

Fig. 4

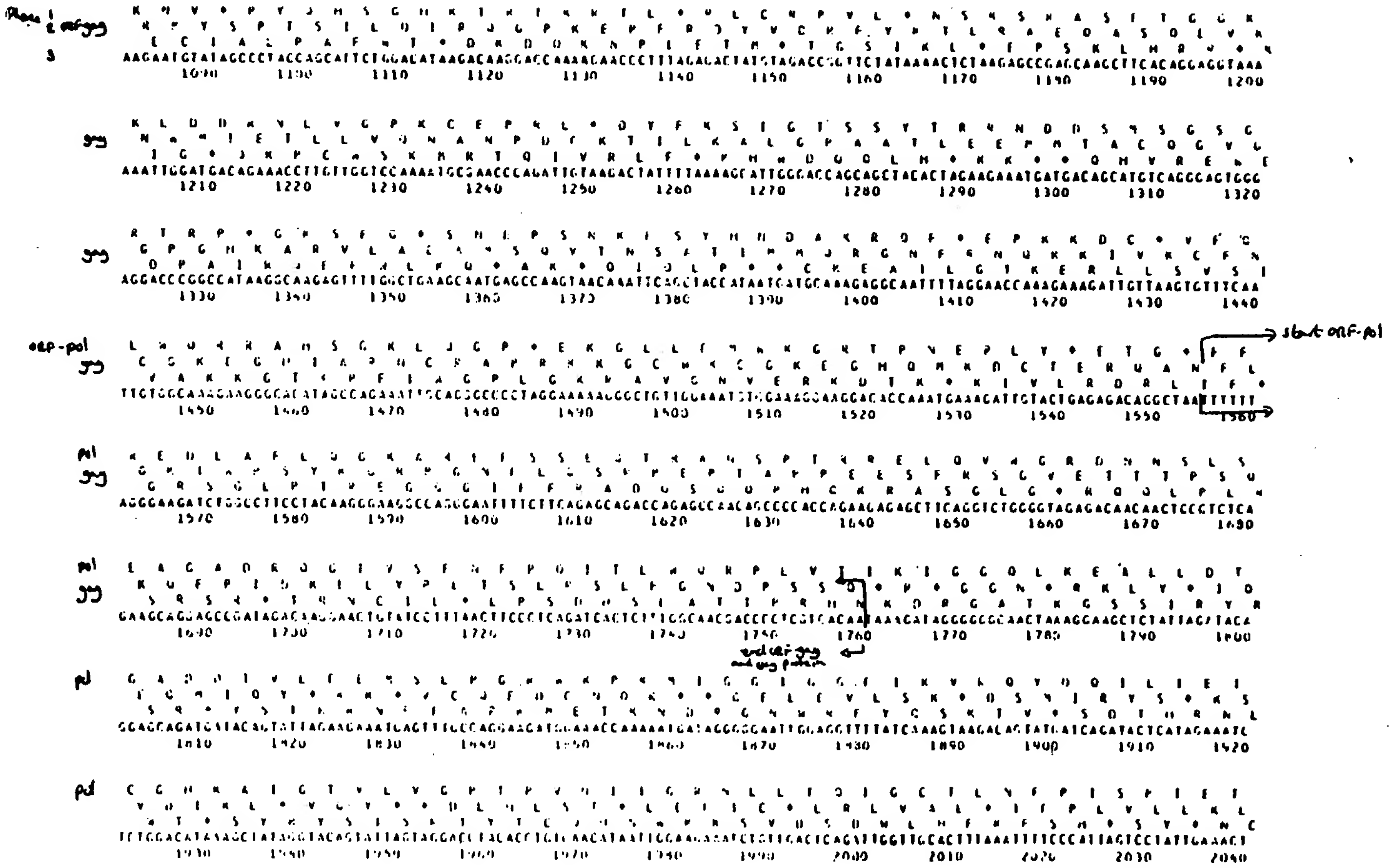

Phase 1 & orf-gag
3
gag
gag
orf-pol
gag
start orf-pol
pol
gag
pol
gag
end orf-gag and gag protein
pol
pol
1090
1100
1110
1120
1130
1140
1150
1160
1170
1180
1190
1200
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
1310
1320
1330
1340
1350
1360
1370
1380
1390
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
1520
1530
1540
1550
1560
1570
1580
1590
1600
1610
1620
1630
1640
1650
1660
1670
1680
1690
1700
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
2030
2040

Fig. 5

771248

08 026736

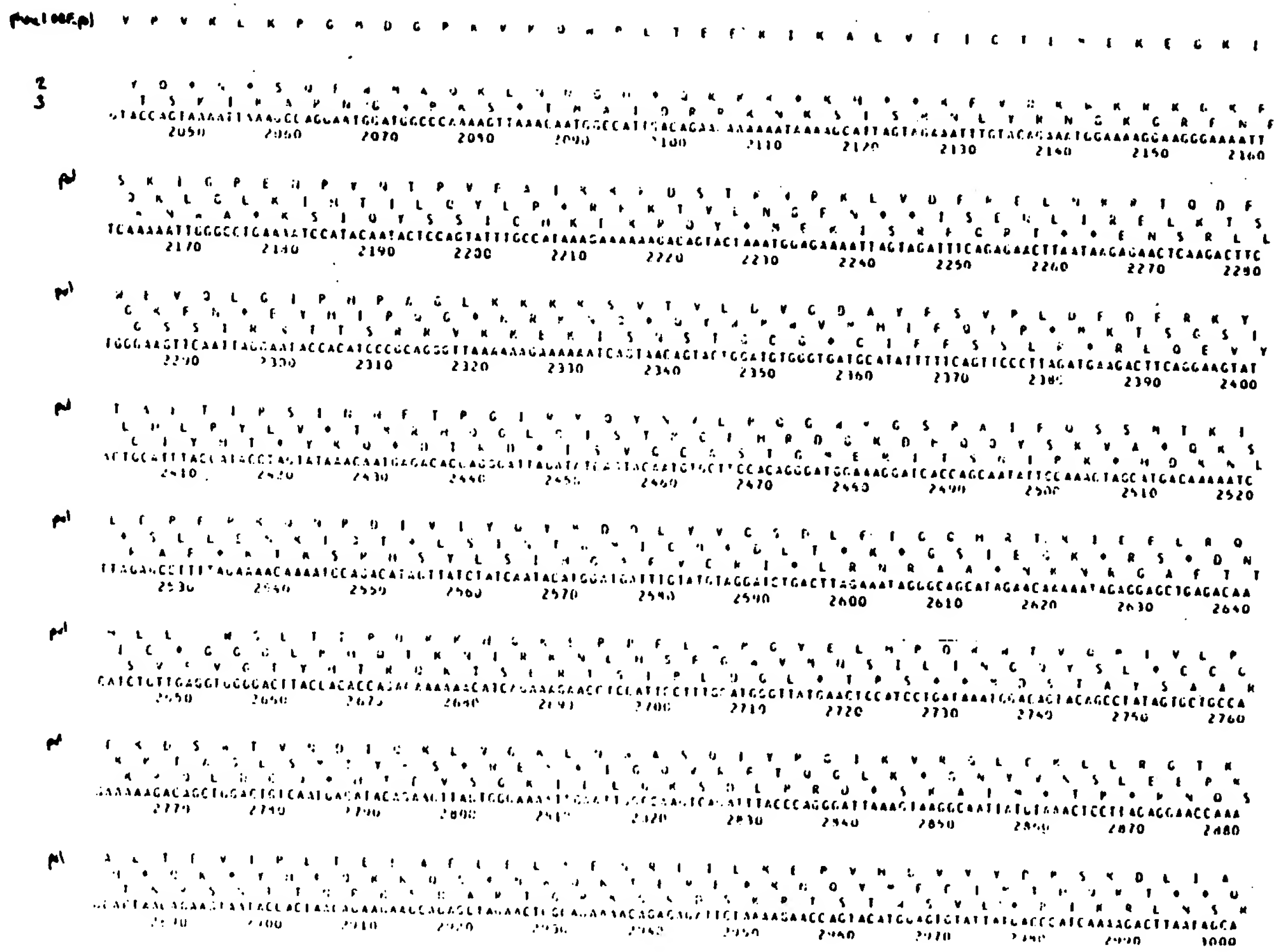

Pmel 008.p)
2
3
2050
2060
2070
2090
2160
2170
2290
2300
2400
2410
2520
2530
2640
2650
2760
2770
2880
2890
3000

Fig 6

28

771248

08 026736

```
ORF-pol  I I Q K Q G D G Q A T Y Q I Y Q I P F K N L P I G K Y A K I N G A H I N D V K Q
          K V R S A G K A N G H I K I I N S H L K I * F Q E N P Q I K G V P T L N N * V N
           N T E A G A P P M D I S N L S K A I * K S F N R * I C K N E G C P H * * C K T I
         GAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAACGAGGGGTGCCCACACTAATGATGTAAAACAA
             3010      3020      3030      3040      3050      3060      3070      3080      3090      3100      3110      3120

Pro1 pol L T E A V Q K I T T E S I V I W G K T P K F K L P I Q K E T W E T W W T E Y W Q

2          * Q K Q C K K * P Q K Y * * Y G E F L L N L N Y P Y K R K H G N H G G Q S I G K
3         N R G S A K N N H R * H S N * G K D S * I * T T H T K G N * G N * V D R V L A S
         TTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTACCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAA
             3130      3140      3150      3160      3170      3180      3190      3200      3210      3220      3230      3240

pol      A T W I P E W E F V N T P P L V K L W Y Q L E K E P I V G A E T F Y V D G A A S
          P P G F L S G S L S I P L L * * N Y G T S * K K N P * * E Q K P S M * W G Q L A
           H L D S * V G V C Q Y P S L S E I M V P V K L R I M S P S * N V L C R W G S * Q
         GCCACCTGGATTCCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACGTTCTATGTAGATGGGGCAGCTAGC
             3250      3260      3270      3280      3290      3300      3310      3320      3330      3340      3350      3360

pol      R E T K L G K A G Y V T N R G R Q K V V T L T D T T N Q K T E L Q A I H L A L Q
          G R L N * E K Q D M L L I E E D K K L S P * L T Q Q I R R L S Y K Q F I * L C R
           G D * I R K S R I C Y * * R K T K S C H P N * H N K S E D * V T S N S S S F A G
         AGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCAG
             3370      3380      3390      3400      3410      3420      3430      3440      3450      3460      3470      3480

pol      D S G L E V N I V T D S Q Y A L G I I Q A Q P D K S E S E L V N Q I I E Q L I K
          I R D * K * I * * Q T H N M H * E S F K H N Q I K V N Q S * S I K * * S S * * K
           F G I R S K Y S N R L T I C I R N H S S T T R * K * I R V S Q S I N R A V N K K
         GATTCGGGATTAGAAGTAAATATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATAAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAA
             3490      3500      3510      3520      3530      3540      3550      3560      3570      3580      3590      3600

pol      K E K V Y L A W V P A H K G I G G N E Q V D K L V S A G I R K V L F L D G I D K
          R K R S I W H G Y Q H T K E L E E M N K * I N * S V L E S G K Y Y F * M E * I N
           G K G L S G M G T S T Q R N W R K * T S R * I S Q C W N Q E S T I F R W N R * G
         AAGGAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTAGATGGAATAGATAAG
             3610      3620      3630      3640      3650      3660      3670      3680      3690      3700      3710      3720

pol      A Q D E H E K Y H S N W R A M A S D F N L P P V V A K E I V A S C D K C Q L K G
          * P K M N M R N I T V I G E Q W L V I L T C H L * * Q K K * * P A V I N V S * K E
           P R * T * E I S Q * L E S N G * * F * P A T C S S N R N S S C L * * M S A K R
         GCCCAAGATGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGA
             3730      3740      3750      3760      3770      3780      3790      3800      3810      3820      3830      3840

pol      E A M H G Q V D C S P G I W Q L D C T H L E G K V I L V A V H V A S G Y I E A E
          K P C M D K * T V V Q E Y G N * I V H I * K E K L S W * Q F M * P V D I * K Q A
           S H A W T S R L * S R N M A T R L Y T F R R K S Y P G S S S C S Q W I Y R S R S
         GAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAA
             3850      3860      3870      3880      3890      3900      3910      3920      3930      3940      3950      3960

pol      V I P A E T G Q E T A Y F L L K L A G R W P V K T I H T D N G S N F T S T T V A
          L F Q Q K Q G R K Q H T F S * N * Q E D G Q * K Q Y I Q T M A A I S P V L * L K
           Y S S R N R A G N S I L S L K I S R K M A S K N N T Y R Q W Q Q F H Q Y Y G * G
         GTTATTCCAGCAGAAACAGGGCAGGAAACAGCATACTTTCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACAGACAATGGCAGCAATTTCACCAGTACTACGGTTAAG
             3970      3980      3990      4000      4010      4020      4030      4040      4050      4060      4070      4080
```

Fig 7

29

11-70

08 026736

```
ORF1 orf-pol  A A C W W A G I K Q E F G I P Y N P Q S Q G V V E S M N K E L K K I I G Q V R D
2              Q P V G G R E S S R N L E F P T I P K V K E • • N L • I K N • R K L • A R • E I
3               S L L V G G N Q A G I W N S L Q S P K S R S S R I Y E • R I K E N Y R P G K R S
              GCAGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGCCAGGTAAGAGAT
                       4090      4100      4110      4120      4130      4140      4150      4160      4170      4180      4190      4200

pol           Q A E H L K T A V Q M A V F I H N F K R K G G I G G Y S A G E R I V D I I A T D

               R L N I L R Q Q Y K W Q Y S S T I L K E K G G L G G T V Q G K E • • T • • Q Q T
                G • T S • D S S T N G S I H P Q F • K K R G D W G V Q C R G K N S R H N S N R H
              CAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGAC
                       4210      4220      4230      4240      4250      4260      4270      4280      4290      4300      4310      4320

pol           I Q T K E L Q K Q I T K I Q N F R V Y Y R D S R D P L W K G P A K L L W K G E G
               Y K L K N Y K N K L Q K F K I F G F I T G T A E I H F G K D Q Q S S S G K V K G
                T N • R I T K T N Y K N S K F S G L L Q G Q Q R S T L E R T S K A P L E R • R G
              ATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGG
                       4330      4340      4350      4360      4370      4380      4390      4400      4410      4420      4430      4440

pol           A V V I Q D N S D I K V V P R R K A K I I R D Y G K Q M A G D D C V A S R Q D E
ORF- Q         Q • • Y K I I V T • K • C Q E E K Q R S L G I M E N R W Q V M I V W Q V D R M R
                S S N T R • • • H K S S A K K K S K D H • G L W K T D G R • • L C G K • T G • G
              GCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAACAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAG
                       4450      4460      4470      4480      4490      4500      4510      4520      4530      4540      4550      4560

pol           D • N M E K F S K T P Y V C F R E S • G M V L • T S L • K P S S K N K F R S T H
Q              I R T W K S L V K H H M Y V S G K A R G W F Y R H H Y E S P H P R I S S E V H I
                L E H G K V • • N T I C M F Q G K L G D G F I D I T M K A L I Q E • V Q K Y T S
              GATTAGAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACAT
                       4570      4580      4590      4600      4610      4620      4630      4640      4650      4660      4670      4680

Q             P T R G C • I G N N N I L G S A Y R R K R L A S G S G S L H R M E E K E I • H T
               P L G D A R L V I T T Y W G L H T G E R D W H L G Q G V S I E W R K K R Y S T Q
                H • G M L D W • • Q H I G V C I Q E K E T G I W V R E S P • N G G K R D I A H K
              CCCACTAGGGGATGCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATCTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACA
                       4690      4700      4710      4720      4730      4740      4750      4760      4770      4780      4790      4800

Q             S R P • T S R P T N S S V L L • L F F R L C Y K K G L I R T Y S • P • V • I S S
               V D P E L A D Q L I H L Y Y F D C F S D S A I R K A L L G H I V S P R C E Y Q A
                • T L N • Q T N • F I C I T L T V F Q T L L • E R P Y • D I • L A L G V N I K Q
              AGTAGACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACATATAGTTAGCCCTAGGTGTGAATATCAAGC
                       4810      4820      4830      4840      4850      4860      4870      4880      4890      4900      4910      4920

Q             R T • Q G R I S T I L G T S S I N N T K K D K A T F A • C Y E T D R G • M E Q A
               G H N K V G S L Q Y L A L A A L I T P K K I K P P L P S V T K L T E D R W N K P
                D I T R • D L Y N T W H • Q H • • H Q K R • S H L C L V L R N • Q R I D G T S P
              AGGACATAACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAATAACACCAAAAAAGATAAAGCCACCTTTGCCTAGTGTTACGAAACTGACAGAGGATAGATGGAACAAGCC
                       4930      4940      4950      4960      4970      4980      4990      5000      5010      5020      5030      5040
                                                                                                                    5031
```

Fig. 8

30

741248

08 026736

```
                                   end ORF-pol
         P E D Q G P Q R E P H N E W T L F L L E E L K N E A V R H F P R I W L H G L G Q
ORF-Q     Q K T K G H R G S H T M N G H * S F * R S L R M K L L D I F L G F G S M A * G N
           R R P K A T E G A T Q * M D T R A F R C A * E * S C * T F S * D L A P W L R A T
         CCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTAGAGCCTTTAGACGAGCTTAAGAATGAAGCTGTTAGACATTTTCCTAGCATTTGGCTCCATGGCTTAGGGCAA
             5050      5060      5070      5080      5090      5100      5110      5120      5130      5140      5150      5160

frame 1  H I Y E I Y G D T W A G V E A I I R I L Q Q L L F I M F P I G C K H S R I G V T
      2   I S N K L M G I L G Q E W K P * * I I C N N C C L S I S E L G V Q I A E * A L L
      3    Y L * N L W G Y L G K S G S H N K N S A I T A V Y P F Q N W V S T * Q N R R Y S
         CATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACT
             5170      5180      5190      5200      5210      5220      5230      5240      5250      5260      5270      5280

         D Q K R A K N G A S P S * T K A L E A S * K S A Q N C L Y H L L L * K V L L S L

         N R G E Q F M E P V D P A L E P W K H P G S Q P K T A C T T C Y C K K C C F H C
          T E E S K K W S Q * I L D * S P G S I Q E V S L K L L V P L A I V K S V A F I A
         CAACAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCACTTGCTATTGTAAAAAGTGTTGCTTTCATTG
             5290      5300      5310      5320      5330      5340      5350      5360      5370      5380      5390      5400
                                         ORF 4
         P S L F H N K S L R H L L W Q E L A E T A T K T S S R Q S D S S S F S I K A V S
          Q V C F T I K A L G I S Y G R K K R R Q R R R P P Q G S G T H Q V S L S K Q * V
           K F V S Q Q K P * A S P M A G R S G D S D E D L L K A V R L I K F L Y Q S S K *
         CCAAGTTTGTTTCACAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGACCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCTATCAAAGCAGTAAGT
             5410      5420      5430      5440      5450      5460      5470      5480      5490      5500      5510      5520

         S T C N A T Y T N S N S S I S S S N N N S N S C V V H S N H R I * E N I K T K K
          V H V M Q P I Q I A I A A L V V A I I I A I V V W S I V I I E Y R K I L R Q R K
           Y M * C N L Y K * Q * Q H * * * Q * * * Q * L C G P * * S * N I G K Y * D K E K
         AGTACATGTAATGCAACCTATACAAATAGCAATAGCAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTAAGACAAAGAAA
             5530      5540      5550      5560      5570      5580      5590      5600      5610      5620      5630      5640

         N R Q V N * * I N K K S R K G W Q * E * K K N I S T C G D G G G W G A P C S L G
          I D K L I D R L I I R A E D S G N E S E G E I S A L V E M G V E M G H H A P W D
ORF-env    * T G * L I D * * K E Q K T V A M R V K E K Y Q H L W R W G W K W G T M L L G I
         AAT..ACAGGTTAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGA
             5650      5660      5670      5680      5690      5700      5710      5720      5730      5740      5750      5760
                            start ORF-env      possible start of env transcription

         Y * * S * V L Q K N C G S W S I W G Y L C G F K Q P P L Y F V H Q N L K H M I Q
          I D D L * C Y R K I V G H S L L W G T C V I G S N H H S I L C I R C * S I * Y F
env        L M I C S A T E K L W V T V Y Y G V P V W K E A T T T L F C A S D A K A Y D T E
         TATTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAG
             5770      5780      5790      5800      5810      5820      5830      5840      5850      5860      5870      5880

         * Y I M F G P H M P V Y P Q T P I H K K * Y W * M * Q K I L T C G A P T W * N R
          G I * C L G H I C L C T H R P G P T R S S I G K C D R K F * H V E K * H G R T D
env        V H N V W A T H A C V P T D P N P Q E V V L V N V T E N F N M W K N D M V E Q M
         AGCTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTAGAACAGA
             5890      5900      5910      5920      5930      5940      5950      5960      5970      5980      5990      6000

         L M K I * S V Y I [illegible] A * S [illegible] V * N * P H S V L V * S A L I W G M L L I P I V V
          [illegible] F * F K V H * F G I C Y * * Y Q * * *
env        [illegible] L * C I D L G N A T N T N S S N
         [illegible]TTAAAGTCCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTA
             6010      6020      6030      6040      6050      6060      6070      6080      6090      6100      6110      6120
```

Fig 9

31

771248

08 026736

```
          I  P  I  V  V  A  G  K  •  •  W  R  K  E  R  •  K  T  A  L  S  I  S  A  Q  A  •  E  V  R  C  R  K  N  M  H  F  F  I  N
           Y  Q  •  •  •  R  G  N  D  D  G  E  R  R  D  K  K  L  L  F  Q  Y  Q  H  K  H  K  R  •  G  A  E  R  I  C  I  F  L  •  T
vpu env     T  N  S  S  S  G  E  M  M  M  E  K  G  E  I  K  N  C  S  F  N  I  S  T  S  I  R  G  K  V  Q  K  E  Y  A  F  F  Y  K  L
         ATACCAATAGTAGTAGCGGGGAAATGATGATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAAC
               6130      6140      6150      6160      6170      6180      6190      6200      6210      6220      6230      6240

fml       L  I  •  Y  Q  •  I  M  I  L  P  A  I  R  •  Q  V  V  T  P  Q  S  L  H  R  P  V  Q  R  Y  P  L  S  Q  F  P  Y  I  I  V
2          •  Y  N  T  N  R  •  •  Y  Y  Q  L  Y  V  D  K  L  •  H  L  S  H  Y  T  G  L  S  K  G  I  L  •  A  N  S  H  T  L  L  C
3 env       D  I  I  P  I  D  N  D  T  T  S  Y  T  L  T  S  C  N  T  S  V  I  T  Q  A  C  P  K  V  S  F  E  P  I  P  I  H  Y  C  A
         TTGATATAATACCAATAGATAATGATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTG
               6250      6260      6270      6280      6290      6300      6310      6320      6330      6340      6350      6360

          P  R  L  V  L  R  F  •  N  V  I  I  R  R  S  M  E  Q  D  H  V  Q  M  S  A  Q  Y  N  V  H  M  E  L  G  Q  •  Y  Q  L  N

           P  G  W  F  C  D  S  K  M  •  •  •  D  V  Q  W  N  R  T  M  Y  K  C  Q  H  S  T  M  Y  T  W  N  •  A  S  S  I  N  S  T
env         P  A  G  F  A  I  L  K  C  N  N  K  T  F  N  G  T  G  P  C  T  N  V  S  T  V  Q  C  T  H  G  I  R  P  V  V  S  T  Q  L
         CCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAAC
               6370      6380      6390      6400      6410      6420      6430      6440      6450      6460      6470      6480

          C  C  •  M  A  V  •  Q  K  K  R  •  •  L  D  L  P  I  S  Q  T  M  L  K  P  •  •  Y  S  •  T  N  L  •  K  L  I  V  Q  D
           A  V  E  W  Q  S  S  R  R  R  G  S  N  •  I  C  Q  F  H  R  Q  C  •  N  H  N  S  T  A  E  P  I  C  R  N  •  L  Y  K  T
env         L  L  N  G  S  L  A  E  E  E  V  V  I  R  S  A  N  F  T  D  N  A  K  T  I  I  V  Q  L  N  Q  S  V  E  I  N  C  T  R  P
         TGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAATCTGTAGAAATTAATTGTACAAGAC
               6490      6500      6510      6520      6530      6540      6550      6560      6570      6580      6590      6600

          P  T  T  I  Q  E  K  V  S  V  S  R  G  D  Q  G  E  H  L  L  Q  •  E  K  •  E  I  •  D  K  H  I  V  T  L  V  E  Q  N  G
           Q  Q  Q  Y  K  K  K  Y  P  Y  P  E  G  T  R  E  S  I  C  Y  N  R  K  N  R  K  Y  E  T  S  T  L  •  H  •  •  S  K  M  E
env         N  N  N  T  R  K  S  I  R  I  Q  R  G  P  G  R  A  F  V  T  I  G  K  I  G  N  M  R  Q  A  H  C  N  I  S  R  A  K  W  N
         CCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGA
               6610      6620      6630      6640      6650      6660      6670      6680      6690      6700      6710      6720

          M  P  L  •  N  R  •  L  A  N  •  E  N  N  L  E  I  I  K  Q  •  S  L  S  N  P  Q  E  G  T  Q  K  L  •  R  T  V  L  I  V
           C  H  F  K  T  D  S  •  Q  I  K  R  T  I  W  K  •  •  N  N  N  L  •  A  I  L  R  R  G  P  R  N  C  N  A  Q  F  •  L  W
env         A  T  L  K  Q  I  A  S  K  L  R  E  Q  F  G  N  N  K  T  I  I  F  K  Q  S  S  G  G  D  P  E  I  V  T  H  S  F  N  C  G
         ATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTG
               6730      6740      6750      6760      6770      6780      6790      6800      6810      6820      6830      6840

          E  G  N  F  S  T  V  I  Q  H  N  C  L  I  V  L  G  L  I  V  L  G  V  L  K  G  Q  I  T  L  K  E  V  T  Q  S  H  S  H  A
           R  G  I  F  L  L  •  F  N  T  T  V  •  •  Y  L  V  •  •  Y  L  E  Y  •  R  V  K  •  H  •  R  K  •  H  N  H  T  P  M  Q
env         G  E  F  F  Y  C  N  S  T  Q  L  F  N  S  T  W  F  N  S  T  W  S  T  E  G  S  N  N  T  E  G  S  D  T  I  T  L  P  C  R
         GAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTCCCATGCA
               6850      6860      6870      6880      6890      6900      6910      6920      6930      6940      6950      6960

          E  •  N  N  L  •  T  C  G  R  K  •  E  K  Q  C  M  P  L  P  S  A  D  K  L  D  V  H  Q  I  L  Q  G  C  Y  •  Q  E  M  V
           N  K  T  I  Y  K  H  V  A  E  S  R  K  S  N  V  C  P  S  H  Q  R  T  N  •  M  F  I  K  Y  Y  R  A  A  I  N  K  R  W  W
env         I  K  Q  F  I  N  M  W  Q  K  V  G  K  A  M  Y  A  P  P  I  S  G  Q  I  R  C  S  S  N  I  T  G  L  L  L  T  R  D  G  G
         GAATAAAACAATTTATAAACATGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTG
               6970      6980      6990      7000      7010      7020      7030      7040      7050      7060      7070      7080

          V  I  A  T  M  S  P  R  S  S  D  L  E  E  E  I  •  G  T  I  G  E  V  N  Y  I  N  I  K  •  •  K  L  N  H  •  E  •  H  P
           •  •  Q  Q  •  V  R  D  L  Q  T  W  R  R  R  Y  E  G  Q  L  E  K  •  I  I  •  I  •  S  S  K  N  •  T  I  R  S  S  T  H
env         N  S  N  N  E  S  E  I  F  R  P  G  G  G  D  M  R  D  N  W  R  S  E  L  Y  K  Y  K  V  V  K  I  E  P  L  G  V  A  P  T
         GTAATAGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCA
               7090      7100      7110      7120      7130      7140      7150      7160      7170      7180      7190      7200
```

Fig 10

771248

08 026736

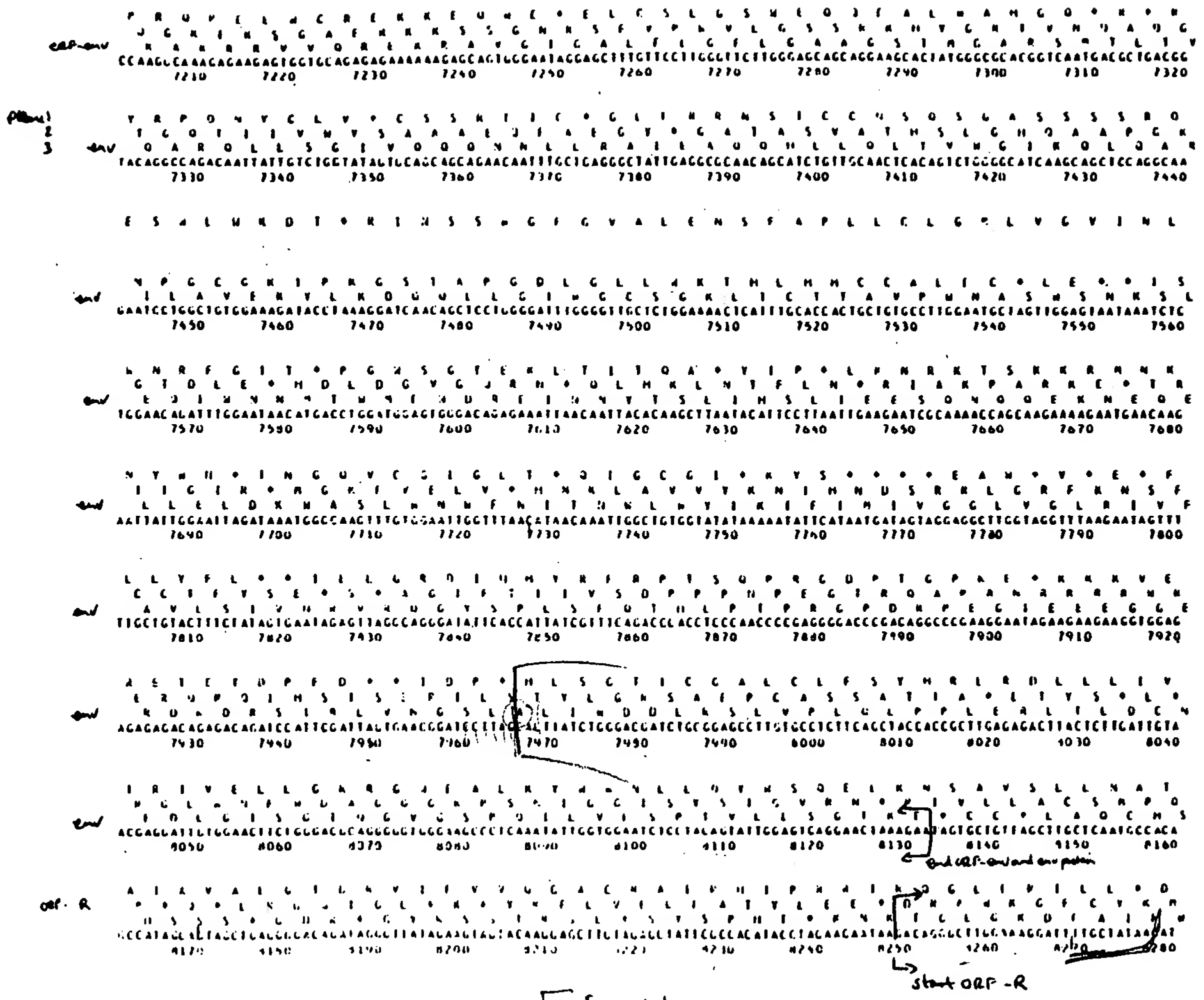

ORF-env
1
2
3
env
ORF-R
7210
7320
7330
7440
7450
7560
7570
7680
7690
7800
7810
7920
7930
8040
8050
8160
8170
8280
End ORF-env and env protein
Start ORF-R

Fig 11

08 026736

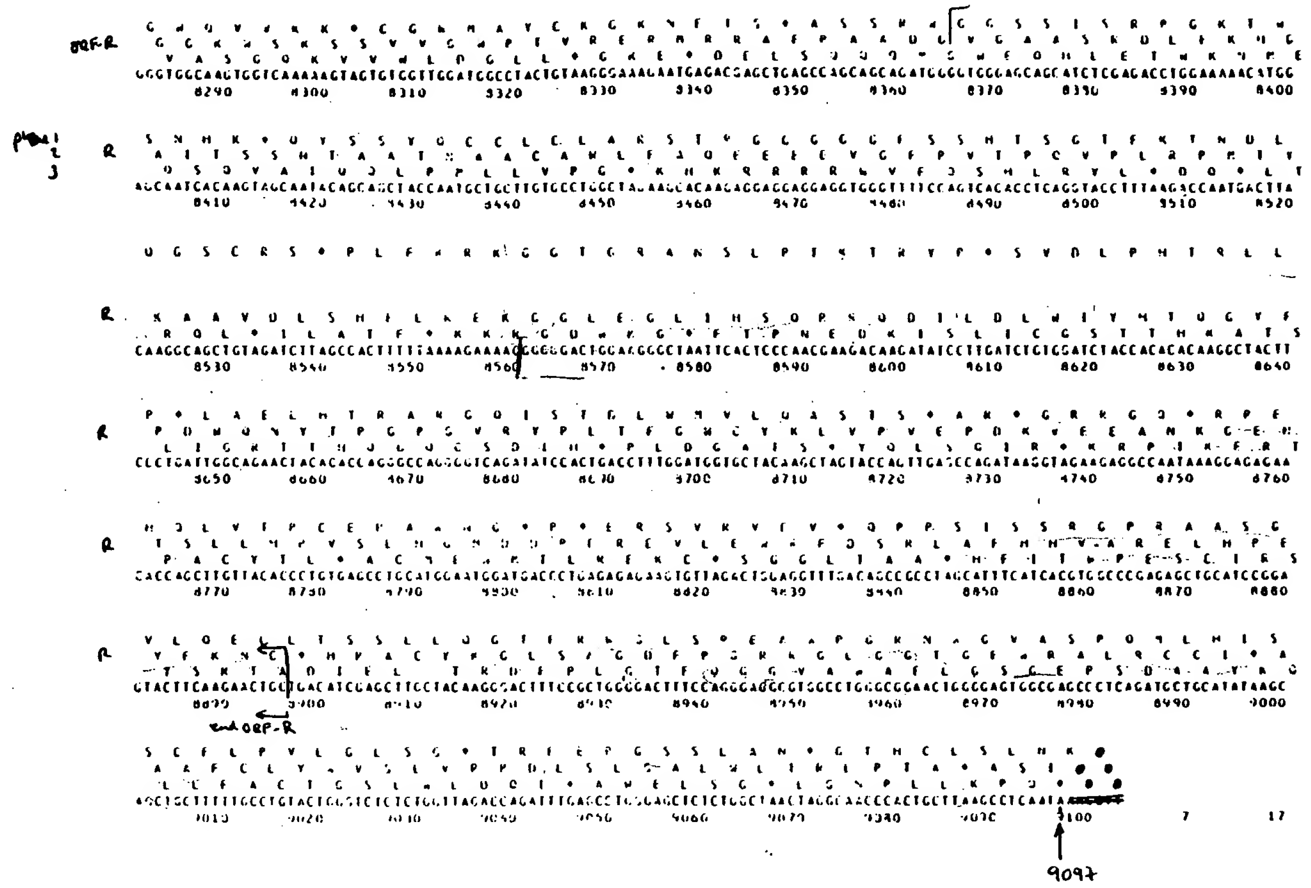

ORF-R
phase 1
2
3
R
end ORF-R
9097
8290
8400
8410
8520
8530
8640
8650
8760
8770
8880
8890
9000
9010
9100

fig 12

34

```
 N  *  G  E  Q  E  *  E  P  V  D  P  R  L  E  P  W  K              P
  T  F  E  S  K  K  W  S  Q  *  I  L  D  *  S  P  G  S  I  Q  E  V  S  L
CAACAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTA
      5290      5300      5310      5320      5330      5340      5350

*  P  S  L  F  H  N  K  S  L  R  H  L  L  W  Q  E  E  A  E  T  A  T  K  T  S
 Q  V  C  F  T  T  K  A  L  G  I  S  Y  G  R  K  K  R  R  Q  R  R  R  P  P
  K  F  V  S  Q  Q  K  P  *  A  S  P  M  A  G  R  S  G  D  S  D  E  D  L
CCAAGTTTGTTTCACAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGACCTCC
      5410      5420      5430      5440      5450      5460      5470

 S  T  C  N  A  T  Y  T  N  S  N  S  S  I  S  S  S  N  N  N  S  N  S  C  V
  V  H  V  M  Q  P  I  Q  I  A  I  A  A  L  V  V  A  I  I  I  A  I  V  V
   Y  M  *  C  N  L  Y  K  *  Q  *  Q  H  *  *  *  Q  *  *  *  Q  *  L  C
AGTACATGTAATGCAACCTATACAAATAGCAATAGCAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTG
      5530      5540      5550      5560      5570      5580      5590

 I  *  Q  V  N  *  *  T  N  R  K  S  R  R  Q  W  Q  *  E  *  R  R  N  I  S
  I  D  R  L  I  D  R  L  I  E  R  A  E  D  S  G  N  E  S  E  G  E  I  S  A
   *  T  G  *  L  I  D  *  *  K  E  Q  K  T  V  A  M  R  V  K  E  K  Y  Q
AATAGACAGGTTAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGC
      5650      5660      5670      5680      5690      5700      5710

 Y  *  *  S  V  V  L  Q  K  N  C  G  S  Q  S  I  M  G  Y  L  C  G  P  K  Q
  I  D  D  L  *  C  Y  R  K  I  V  G  H  S  L  L  W  G  T  C  V  E  G  S  N
   L  M  I  C  S  A  T  E  K  L  W  V  T  V  Y  Y  G  V  P  V  W  K  E  A
TATTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAA
      5770      5780      5790      5800      5810      5820      5830

 K  Y  I  M  F  G  P  H  M  P  V  Y  P  Q  T  P  T  H  K  K  *  Y  *  *  *
  G  T  *  C  L  G  H  T  C  L  C  T  H  R  P  Q  P  T  R  S  S  I  G  *
   V  H  N  V  W  A  T  H  A  C  V  P  T  D  P  N  P  Q  E  V  V  L  V  N
AGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATG
      5890      5900      5910      5920      5930      5940      5950

 C  M  R  I  *  S  V  Y  G  I  K  A  *  S  H  V  *  N  *  P  H  S  V  L  V
  A  *  G  Y  N  Q  F  M  G  S  K  P  K  A  M  C  K  I  N  P  T  L  C  *  F
   H  E  D  I  I  S  L  W  D  Q  S  L  K  P  C  V  K  L  T  P  L  C  V  S
TGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTAGTT
      6010      6020      6030      6040      6050      6060      6070

 I  P  I  V  V  A  G  K  *  *  W  R  K  E  R  *  K  T  A  L  S  I  S  A  Q
  Y  Q  *  *  *  K  G  N  D  D  G  E  R  R  D  K  K  L  L  F  Q  Y  Q  H  K
   T  N  S  S  S  G  E  M  M  M  E  K  G  E  I  K  N  C  S  F  N  I  S  T
ATACCAATAGTAGTAGCGGGGAAATGATGATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAA
      6130      6140      6150      6160      6170      6180      6190

 L  I  *  Y  Q  *  I  M  I  L  P  A  I  R  *  Q  V  V  T  P  Q  S  L  H  R
  *  Y  N  T  N  R  *  *  Y  Y  Q  L  Y  V  D  K  L  *  H  L  S  H  Y  T  G
   D  I  I  P  I  D  N  D  T  T  S  Y  T  L  T  S  C  N  T  S  V  I  T  Q
TTGATATAATACCAATAGATAATGATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGG
      6250      6260      6270      6280      6290      6300      6310

 P  R  L  V  L  R  F  *  N  V  I  I  R  R  S  M  E  Q  D  H  V  Q  *  S  A
```

Fig 14

```
 G  S  Q  P  K  T  A  C  T  T  C  Y  C  K  K  C  C  F  H  C
Q  E  V  S  L  K  L  L  V  P  L  A  I  V  K  S  V  A  F  I  A
AGGAAGTCAGCCTAAAACTGCTTGTACCACTTGCTATTGTAAAAAGTGTTGCTTTCATTG
      5350      5360      5370      5380      5390      5400

 A  T  K  T  S  S  R  Q  S  D  S  S  S  F  S  I  K  A  V  S
 R  R  R  P  P  Q  G  S  C  T  H  C  V  S  L  S  K  Q  *  V
S  D  E  D  L  L  K  A  V  R  L  I  K  F  L  Y  Q  S  S  K  *
GCGACGAAGACCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCTATCAAAGCAGTAAGT
      5470      5480      5490      5500      5510      5520

 S  N  S  C  V  V  H  S  N  H  R  I  *  E  N  I  K  T  K  K
 A  I  V  V  W  S  I  V  I  I  E  Y  R  K  I  L  R  Q  R  K
*  Q  *  L  C  G  P  *  *  S  *  N  I  G  K  Y  *  D  K  E  K
AGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTAAGACAAAGAAA
      5590      5600      5610      5620      5630      5640

 R  R  N  I  S  T  C  G  D  G  G  G  N  G  A  P  C  S  L  G
  G  E  I  S  A  L  V  E  M  G  V  E  M  G  H  H  A  P  W  D
K  E  K  Y  Q  H  L  W  R  W  G  W  K  W  G  T  M  L  L  G  I
AGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGA
      5710      5720      5730      5740      5750      5760

 C  G  F  K  Q  P  P  L  Y  F  V  H  Q  M  L  K  H  M  I  Q
  V  E  G  S  N  H  H  S  I  L  C  I  Q  C  *  S  I  *  Y  R
V  W  K  E  A  T  T  T  L  F  C  A  S  D  A  K  A  Y  D  T  E
TGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAG
      5830      5840      5850      5860      5870      5880

 *  Y  *  *  M  *  Q  K  I  L  T  C  G  K  M  T  W  *  N  R
  S  I  G  K  C  D  R  K  F  *  H  V  E  K  *  H  G  R  T  D
V  V  L  V  N  V  T  E  N  F  N  M  W  K  N  D  M  V  E  Q  M
TAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTAGAACAGA
      5950      5960      5970      5980      5990      6000

 H  S  V  L  V  *  S  A  L  T  W  G  *  L  L  I  P  I  V  V
  T  L  C  *  F  K  V  H  *  F  G  E  C  Y  *  Y  Q  *  *  *
Q  L  C  V  S  L  K  C  T  D  L  G  N  A  T  N  T  N  S  S  N
CACTCTGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTA
      6070      6090      6090      6100      6110      6120

 S  I  S  A  Q  A  *  E  V  R  C  P  K  N  M  H  F  F  I  N
  Q  Y  Q  H  K  H  K  R  *  G  A  E  R  I  C  I  F  L  *  T
F  N  I  S  T  S  I  R  G  K  V  Q  K  E  Y  A  F  F  Y  K  L
TCAATATCACCACAAGCATAAGAGGTAAGGTCCAGAAAGAATATGCATTTTTTTATAAAC
      6190      6200      6210      6220      6230      6240

 Q  S  L  H  R  P  V  Q  R  Y  P  L  S  Q  F  P  Y  I  I  V
  S  H  Y  T  G  L  S  K  G  I  L  *  A  N  S  H  T  L  L  C
S  V  I  T  Q  A  C  P  K  V  S  F  E  P  I  P  I  H  Y  C  A
CAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTG
      6310      6320      6330      6340      6350      6360

V  Q  M  S  A  Q  Y  N  V  H  M  F  L  G  G  *  Y  Q  L  N
```

Fig 15

```
  P  G  W  F  C  D  S  K  I  *  |  *  *  D  V  Q  W  N  R  T  M  Y  K  C  Q
   P  A  G  F  A  I  L  K  S  N  N  K  T  F  N  G  T  G  P  C  T  N  V  S
CCCCGGCTGGTTTTGCGATTCTAAAATCTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGC
     6370      6380      6390      6400      6410      6420      6430

  C  C  *  M  A  V  *  Q  K  K  R  *  *  L  D  L  P  I  S  Q  T  M  L  K  P
   A  V  E  W  Q  S  S  R  R  R  G  S  N  *  I  C  Q  F  H  R  Q  C  *  N
  L  L  N  G  S  L  A  E  E  E  V  V  I  R  S  A  N  F  T  D  N  A  K  T
TGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACC
     6490      6500      6510      6520      6530      6540      6550

  P  T  T  I  Q  E  K  V  S  V  S  R  G  D  Q  G  E  H  L  L  Q  *  E  K  *
   Q  Q  Q  Y  K  K  K  Y  P  Y  P  E  G  T  R  E  S  I  C  Y  N  R  K  N
  N  N  N  T  R  K  S  I  R  I  Q  R  G  P  G  R  A  F  V  T  I  G  K  I
CCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATA
     6610      6620      6630      6640      6650      6660      6670

  M  P  L  *  N  R  *  L  A  N  *  E  N  N  L  E  I  I  K  Q  *  S  L  S  N
   C  H  F  K  T  D  S  *  Q  I  K  R  T  I  W  K  *  *  N  N  N  L  *  A
  A  T  L  K  Q  I  A  S  K  L  R  E  Q  F  G  N  N  K  T  I  I  F  K  Q
ATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAA
     6730      6740      6750      6760      6770      6780      6790

  I  G  N  F  S  T  V  I  Q  H  N  C  L  I  V  L  G  L  I  V  L  G  V  L  K
   R  G  I  F  L  L  *  F  N  T  T  V  *  *  Y  L  V  *  *  Y  L  E  Y  *
  G  E  F  F  Y  C  N  S  T  Q  L  F  N  S  T  W  F  N  S  T  W  S  T  E
GAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAA
     6850      6860      6870      6880      6890      6900      6910

  E  *  N  N  L  *  T  C  G  R  K  *  E  K  Q  C  M  P  L  P  S  A  D  K  L
   N  K  T  I  Y  K  H  V  A  G  S  R  K  S  N  V  C  P  S  H  Q  R  T  N
  I  K  Q  F  I  N  M  W  Q  E  V  G  K  A  M  Y  A  P  P  I  S  G  Q  I
GAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGCGGACAAATT
     6970      6980      6990      7000      7010      7020      7030

  V  I  T  T  M  G  P  R  S  S  D  L  E  E  E  I  *  G  T  I  G  E  V  N  Y
   *  *  Q  Q  W  V  R  D  L  Q  T  W  R  R  R  Y  E  G  Q  L  E  K  *  I  I
  N  N  N  N  G  S  E  I  F  R  P  G  G  G  D  M  R  D  N  W  R  S  E  L
GTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTAT
     7090      7100      7110      7120      7130      7140      7150

  P  R  Q  R  E  E  W  C  R  E  K  K  E  Q  W  E  *  E  L  C  S  L  G  S  W
   Q  G  K  E  K  S  G  A  E  R  K  K  S  S  G  N  R  S  F  V  P  W  V  L  G
  K  A  K  R  R  V  V  Q  R  E  K  R  A  V  G  I  G  A  L  F  L  G  F  L
CCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGG
     7210      7220      7230      7240      7250      7260      7270

  Y  R  P  D  N  Y  C  L  V  *  C  S  S  R  T  I  C  *  G  L  L  R  R  N  S
   T  G  Q  T  I  I  V  W  Y  S  A  A  A  E  Q  F  A  E  G  Y  *  G  A  T  A
  Q  A  R  Q  L  L  S  G  I  V  Q  Q  Q  N  N  L  L  R  A  I  E  A  Q  Q
TACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGC
     7330      7340      7350      7360      7370      7380      7390

  E  S  A  L  W  K  D  T  *  R  I  N  S  S  W  G  F  G  V  A  L  E  N  S  F
```

37

Fig. 16

771248

08 026736

```
 N   R   T   M   Y   K   C   Q   H   S   T   M   Y   T   W   N   *   A   S   S   I   N   S   T
   T   G   P   C   T   N   V   S   T   V   Q   C   T   H   G   I   R   *   V   V   S   T   Q   L
AACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAAC
      6420      6430      6440      6450      6460      6470      6490

 P   I   S   Q   T   M   L   K   P   *   *   Y   S   *   T   N   L   *   K   L   I   V   Q   D
  Q   F   H   R   Q   C   *   N   H   N   S   T   A   E   P   I   C   R   N   *   L   Y   K   T
   N   F   T   D   N   A   K   T   I   I   V   Q   L   N   Q   S   V   E   I   N   C   T   R   P
CAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAATCTGTAGAAATTAATTGTACAAGAC
      6540      6550      6560      6570      5580      6590      6600

 F   H   L   L   Q   *   E   K   *   E   I   *   D   K   H   I   V   T   L   V   F   Q   N   G
  S   I   C   Y   N   *   K   N   *   K   Y   E   T   S   T   L   *   H   *   *   S   K   M   E
   A   F   V   T   I   G   K   I   G   N   *   R   Q   A   H   C   N   I   S   R   A   K   W   N
AGCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGA
      6660      6670      6680      6690      6700      6710      6720

 I   I   K   Q   *   S   L   S   N   P   Q   E   G   T   Q   K   L   *   R   T   V   L   I   V
  *   *   N   N   N   L   *   A   I   L   R   R   G   P   R   N   C   N   A   Q   F   *   L   W
   N   K   T   I   I   F   K   Q   S   S   G   G   D   P   E   I   V   T   H   S   F   N   C   G
TAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTG
      6780      6790      6800      6810      6920      6830      6840

 L   I   V   L   G   V   L   K   G   Q   I   T   L   K   E   V   T   Q   S   H   S   H   A
  V   *   *   Y   L   E   Y   *   R   V   K   *   H   *   R   K   *   H   V   H   T   P   M   C
   F   N   S   T   W   S   T   E   G   S   N   N   T   E   G   S   D   T   I   T   L   P   C   R
TTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTCCCATGCA
      6900      6910      6920      6930      6940      6950      6960

 *   P   L   P   S   A   D   K   L   D   V   H   Q   I   L   Q   G   C   Y   *   Q   E   M   V
  C   P   S   H   Q   R   T   N   *   M   F   I   K   Y   Y   R   A   A   I   N   K   R   W   W
   A   P   P   I   S   G   Q   I   R   C   S   S   N   I   T   G   L   L   L   T   R   D   G   G
TGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTG
      7020      7030      7040      7050      7060      7070      7080

 *   G   T   I   G   E   V   N   Y   I   N   I   K   *   *   K   L   N   H   *   E   *   H   P
  E   G   Q   L   E   K   *   I   I   *   I   *   S   S   K   N   *   T   I   R   S   S   T   H
   R   D   N   W   R   S   E   L   Y   K   Y   K   V   V   K   I   E   P   L   G   V   A   P   T
GAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCA
      7140      7150      7160      7170      7180      7190      7200

 *   E   L   C   S   L   G   S   W   E   Q   Q   E   A   L   *   A   H   G   Q   *   R   *   R
  R   S   F   V   P   W   V   L   G   S   S   R   K   H   Y   G   R   T   V   N   D   A   D   G
   G   A   L   F   L   G   F   L   G   A   A   G   S   T   M   G   A   R   S   M   T   L   T   V
AGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCACGGTCAATGACGCTGACGG
      7260      7270      7290      7290      7300      7310      7320

 *   *   G   L   L   R   R   N   S   I   C   C   N   S   Q   S   G   A   S   S   S   S   R   Q
  A   E   G   Y   *   G   A   T   A   S   V   A   T   H   S   L   G   H   Q   A   A   P   G   K
   L   R   A   I   E   A   Q   Q   H   L   L   Q   L   T   V   W   G   I   K   Q   L   Q   A   R
GCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAA
      7380      7390      7400      7410      7420      7430      7440

  G   V   A   L   F   N   S   F   A   P   L   L   C   L   G   *   L   V   G   V   I   N   L
```

36

Fig 17

771248

08 026736

```
  N   P   G   C   G   K   I   P   K   G   S   T   A   P   G   D   L   G   L   L
    I   L   A   V   E   R   Y   L   K   D   Q   Q   L   L   G   I   W   G   C   S   G   K   L   I
GAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCAT
      7450      7460      7470      7480      7490      7500      7510

  W   N   R   F   G   I   T   *   P   G   W   S   G   T   E   K   L   T   I   T   Q   A   *   Y
    G   T   D   L   E   *   H   D   L   D   G   V   G   Q   R   N   *   Q   L   H   K   L   N   T
      E   Q   I   W   N   N   M   T   W   M   E   W   D   R   E   I   N   N   Y   T   S   L   I   H
TGGAACAGATTTGGAATAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACA
      7570      7580      7590      7600      7610      7620      7630

  N   Y   *   N   *   I   N   G   Q   V   C   Q   I   G   L   T   *   Q   I   G   C   G   I   *
    I   I   G   I   R   *   M   G   K   F   V   E   L   V   *   H   N   K   L   A   V   V   Y   K
      L   L   E   L   D   K   W   A   S   L   W   N   W   F   N   I   T   N   W   L   W   Y   I   K
AATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAA
      7690      7700      7710      7720      7730      7740      7750

  L   L   Y   F   L   *   *   I   E   L   G   R   D   I   H   H   Y   R   F   R   P   T   S   Q
    C   C   T   F   Y   S   E   *   S   *   A   G   I   F   T   I   I   V   S   D   P   P   P   N
      A   V   L   S   I   V  /N   R   V   R   Q   G   Y   S   P   L   S   F   Q   T   H   L   P   T
TTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAAC
      7810      7820      7830      7840      7850      7860      7870

  R   E   T   E   T   D   P   F   D   *   *   T   D   P   *   H   L   S   G   T   I   C   G   A
    E   R   Q   R   Q   I   H   S   I   S   E   R   I   L   S   T   Y   L   G   R   S   A   E   P
      R   D   R   D   R   S   I   R   L   V   N   G   S   L   A   L   I   W   D   D   L   R   S   L
AGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTGCGGAGCCT
      7930      7940      7950      7960      7970      7980      7990

  T   R   I   V   E   L   L   G   R   R   G   W   E   A   L   K   Y   W   W   N   L   L   Q   Y
    R   G   L   W   N   F   W   D   A   G   G   G   K   P   S   N   I   G   G   I   S   Y   S   I
      E   D   C   G   T   S   G   T   Q   G   V   G   S   P   Q   I   L   V   E   S   P   T   V   L
ACGAGGATTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAATCTCCTACAGTATT
      8050      8060      8070      8080      8090      8100      8110

  A   I   A   V   A   E   G   T   D   R   V   I   E   V   V   Q   G   A   C   R   A   I   R   H   I
    P   *   Q   *   L   R   G   Q   I   G   L   *   K   *   Y   K   E   L   V   E   L   F   A   T
      H   S   S   S   *   G   D   R   *   G   Y   R   S   S   T   R   S   L   *   S   Y   S   P   H
GCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACAT
      8170      8180      8190      8200      8210      8220      8230

  G   W   Q   V   V   K   K   *   C   G   W   M   A   Y   C   K   G   K   N   E   T   S   *   A   S
    G   G   K   W   S   K   S   S   V   V   G   W   P   T   V   R   E   R   M   R   R   A   E   P
      V   A   S   G   Q   K   V   V   W   L   D   G   L   L   *   G   K   E   *   D   E   L   S   Q
GGGTGGCAAGTGGTCAAAAAGTAGTGTGGTTGGATGGCCTACTGTAAGGGAAAGAATGAGACGAGCTGAGCCAG
      8290      8300      8310      8320      8330      8340      8350

  S   N   H   K   *   Q   Y   S   S   Y   Q   C   C   L   C   L   A   R   S   T   R   G   G   G   G
    A   I   T   S   S   N   T   A   A   T   N   A   A   C   A   W   L   F   A   Q   E   E   E   E
      Q   S   Q   V   A   I   Q   Q   L   P   M   L   L   V   P   G   *   K   H   K   R   R   R
AGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAAGAGGAGCAGGAGC
      8410      8420      8430      8440      8450      8460      8470

  Q   G   S   C   R   S   *   P   L   F   K   R   K   G   G   T   G   R   A   N   S   L   P   T
```

39

15

Fig 18

08 026736

```
  K   T   H   L   H   H   C   C   A   L   E   C   *   L   E   *   *   I   S
   G   K   L   I   C   T   T   A   V   P   W   N   A   S   W   S   N   K   S   L
TGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTC
       7510      7520      7530      7540      7550      7560

Q   A   *   Y   I   P   *   L   K   N   R   K   T   S   K   K   R   M   N   K
  K   L   N   T   F   L   N   *   R   I   A   K   P   A   R   K   E   *   T   R
)   S   L   I   H   S   L   I   E   E   S   Q   N   Q   Q   E   K   N   E   Q   E
AAGCTTAATACATTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAG
       7630      7640      7650      7660      7670      7680

C   G   I   *   K   Y   S   *   *   *   *   E   A   W   *   V   *   E   *   F
  V   V   Y   K   N   I   H   N   D   S   R   R   L   G   R   F   K   N   S   F
    W   Y   I   K   I   F   I   M   I   V   G   G   L   V   G   L/  R/  I   V   F
GTGGTATATAAAAATATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTT
       7750      7760      7770      7780      7790      7800

    T   S   Q   P   R   G   D   P   T   G   P   K   E   *   K   K   K   V   E
  P   P   P   N   P   E   G   T   R   Q   A   R   R   N   R   R   R   R   W   R
    H   L   P   T   P   R   G   P   D   R   P   E   G   I   E   E   E   G   G   E
CCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAG
       7870      7880      7890      7900      7910      7920

I   C   G   A   L   C   L   F   S   Y   H   R   L   R   D   L   L   L   I   V
  S   A   E   P   C   A   S   S   A   T   T   A   *   E   T   Y   S   *   L   *
    L   R   S   L   V   P   L   Q   L   P   P   L   E   R   L   T   L   D   C   N
TCTGCGGAGCCTTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTA
       7990      8000      8010      8020      8030      8040

L   L   Q   Y   W   S   Q   E   L   K   N   S   A   V   S   L   L   N   A   T
  S   Y   S   I   G   V   R   N   *   R   I   V   L   L   A   C   S   M   P   Q
    P   T   V   L   E   S   G   T   K   E   *   C   C   *   L   A   Q   C   H   S
TCCTACAGTATTGGAGTCAGGAACTAAAGAATAGTGCTGTTAGCTTGCTCAATGCCACA
       8110      8120      8130      8140      8150      8160

    I   R   H   I   P   R   R   I   R   Q   G   L   E   R   I   L   L   *   D
  L   F   A   T   Y   L   E   E   *   D   R   A   W   K   G   F   C   Y   K   M
    Y   S   P   H   T   *   K   N   K   T   G   L   G   K   D   F   A   I   R   W
TATTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGATTTTGCTATAAGAT
       8230      8240      8250      8260      8270      8280

T   S   *   A   S   S   R   W   G   G   S   S   I   S   R   P   G   K   T   W
  R   A   E   P   A   A   D   G   V   G   A   A   S   R   D   L   E   K   H   G
    E   L   S   Q   Q   Q   *   G   W   E   Q   H   L   E   T   W   K   N   M   E
GAGCTGAGCCAGCAGCAGATGGGGTGGAGCAGCATCTCGAGACCTGGAAAAACATGG
       8350      8360      8370      8380      8390      8400

    G   G   G   G   G   F   S   S   H   T   S   G   T   F   K   T   N   D   L
  E   E   E   E   V   G   F   P   V   T   P   Q   V   P   L   R   P   M   T   Y
    R   R   R   R   W   V   F   Q   S   H   L   R   Y   L   *   D   Q   *   L   T
AGGAGCAGGAGGGGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTA
                 8480      8490      8500      8510      8520

  L   P   T   K   T   R   Y   P   *   S   V   D   L   P   H   T   R   L   L
```

Junie track

15/15 B/L

Fig 19

```
        10         20         30         40         50         60
AAGCTTGCCT TGAGTGCTTC AAGTAGTGTG TGCCCGTCTG TTGTGTGACT CTGGTAACTA

        70         80         90        100        110        120
GAGATCCCTC AGACCCTTTT AGTCAGTGTG GAAAATCTCT AGCAGTGGCG CCCGAACAGG

       130        140        150        160        170        180
GACTTGAAAG CGAAAGGGAA ACCAGAGGAG CTCTCTCGAC GCAGGACTCG GCTTGCTGAA

       190        200        210        220        230        240
GCGCGCACGG CAAGAGGCGA GGGGAGGCGA CTGGTGAGTA CGCCAAAAAT TTTGACTAGC

       250        260        270        280        290        300
GGAGGCTAGA AGGAGAGAGA TGGGTGCGAG AGCGTCAGTA TTAAGCGGGG GAGAATTAGA

       310        320        330        340        350        360
TCGATCGGAA AAAATTCGGT TAAGGCCAGG GGGAAAGAAA AAATATAAAT TAAAACATAT

       370        380        390        400        410        420
AGTATGGGCA AGCAGGGAGC TAGAACGATT CGCTGTTAAT CCTGGCCTGT TAGAAACATC

       430        440        450        460        470        480
AGAAGGCTGT AGACAAATAC TGGGACAGCT ACAACCATCC CTTCAGACAG GATCAGAAGA

       490        500        510        520        530        540
ACTTAGATCA TTATATAATA CAGTAGCAAC CCTCTATTGT GTGCATCAAA GGATAGAGAT

       550        560        570        580        590        600
AAAAGACACC AAGGAAGCTT TAGACAAGAT AGAGGAAGAG CAAAACAAAA GTAAGAAAAA

       610        620        630        640        650        660
AGCACAGCAA GCAGCAGCTG ACACAGGACA CAGCAGCCAG GTCAGCCAAA ATTACCCTAT

       670        680        690        700        710        720
AGTGCAGAAC ATCCAGGGGC AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAATGC

       730        740        750        760        770        780
ATGGGTAAAA GTAGTAGAAG AGAAGGCTTT CAGCCCAGAA GTGATACCCA TGTTTTCAGC

       790        800        810        820        830        840
ATTATCAGAA GGAGCCACCC CACAAGATTT AAACACCATG CTAAACACAG TGGGGGGACA

       850        860        870        880        890        900
TCAAGCAGCC ATGCAAATGT TAAAAGAGAC CATCAATGAG GAAGCTGCAG AATGGGATAG

       910        920        930        940        950        960
AGTGCATCCA GTGCATGCAG GGCCTATTGC ACCAGGCCAG ATGAGAGAAC CAAGGGGAAG

       970        980        990       1000       1010       1020
TGACATAGCA GGAACTACTA GTACCCTTCA GGAACAAATA GGATGGATGA CAAATAATCC

      1030       1040       1050       1060       1070       1080
ACCTATCCCA GTAGGAGAAA TTTATAAAAG ATGGATAATC CTGGGATTAA ATAAAATAGT

      1090       1100       1110       1120       1130       1140
```

Fig 20

```
AAGAATGTAT AGCCCTACCA GCATTCTGGA CATAAGACAA GGACCAAAAG AACCCTTTAG

      1150       1160       1170       1180       1190       1200
AGACTATGTA GACCGGTTCT ATAAAACTCT AAGAGCCGAG CAAGCTTCAC AGGAGGTAAA

      1210       1220       1230       1240       1250       1260
AAATTGGATG ACAGAAACCT TGTTGGTCCA AAATGCGAAC CCAGATTGTA AGACTATTTT

      1270       1280       1290       1300       1310       1320
AAAAGCATTG GGACCAGCAG CTACACTAGA AGAAATGATG ACAGCATGTC AGGGAGTGGG

      1330       1340       1350       1360       1370       1380
AGGACCCGGC CATAAGGCAA GAGTTTTGGC TGAAGCAATG AGCCAAGTAA CAAATTCAGC

      1390       1400       1410       1420       1430       1440
TACCATAATG ATGCAAAGAG GCAATTTTAG GAACCAAAGA AAGATTGTTA AGTGTTTCAA

      1450       1460       1470       1480       1490       1500
TTGTGGCAAA GAAGGGCACA TAGCCAGAAA TTGCAGGGCC CCTAGGAAAA AGGGCTGTTG

      1510       1520       1530       1540       1550       1560
GAAATGTGGA AAGGAAGGAC ACCAAATGAA AGATTGTACT GAGAGACAGG CTAATTTTTT

      1570       1580       1590       1600       1610       1620
AGGGAAGATC TGGCCTTCCT ACAAGGGAAG GCCAGGGAAT TTTCTTCAGA GCAGACCAGA

      1630       1640       1650       1660       1670       1680
GCCAACAGCC CCACCAGAAG AGAGCTTCAG GTCTGGGGTA GAGACAACAA CTCCCTCTCA

      1690       1700       1710       1720       1730       1740
GAAGCAGGAG CCGATAGACA AGGAACTGTA TCCTTTAACT TCCCTCAGAT CACTCTTTGG

      1750       1760       1770       1780       1790       1800
CAACGACCCC TCGTCACAAT AAAGATAGGG GGGCAACTAA AGGAAGCTCT ATTAGATACA

      1810       1820       1830       1840       1850       1860
GGAGCAGATG ATACAGTATT AGAAGAAATG AGTTTGCCAG GAAGATGGAA ACCAAAAATG

      1870       1880       1890       1900       1910       1920
ATAGGGGGAA TTGGAGGTTT TATCAAAGTA AGACAGTATG ATCAGATACT CATAGAAATC

      1930       1940       1950       1960       1970       1980
TGTGGACATA AAGCTATAGG TACAGTATTA GTAGGACCTA CACCTGTCAA CATAATTGGA

      1990       2000       2010       2020       2030       2040
AGAAATCTGT TGACTCAGAT TGGTTGCACT TTAAATTTTC CCATTAGTCC TATTGAAACT

      2050       2060       2070       2080       2090       2100
GTACCAGTAA AATTAAAGCC AGGAATGGAT GGCCCAAAAG TTAAACAATG GCCATTGACA

      2110       2120       2130       2140       2150       2160
GAAGAAAAAA TAAAAGCATT AGTAGAAATT TGTACAGAAA TGGAAAAGGA AGGGAAAATT

      2170       2180       2190       2200       2210       2220
TCAAAAATTG GGCCTGAAAA TCCATACAAT ACTCCAGTAT TTGCCATAAA GAAAAAAGAC

      2230       2240       2250       2260       2270       2280
AGTACTAAAT GGAGAAAATT AGTAGATTTC AGAGAACTTA ATAAGAGAAC TCAAGACTTC

      2290       2300       2310       2320       2330       2340
TGGGAAGTTC AATTAGGAAT ACCACATCCC GCAGGGTTAA AAAAGAAAAA ATCAGTAACA

      2350       2360       2370       2380       2390       2400
```

Fig 21

```
[illegible] TGGGTGATGC ATATTTTTCA GTTCCCTTAG ATGAAGACTT CAGGAAGTAT

      2410       2420       2430       2440       2450       2460
ACTGCATTTA CCATACCTAG TATAAACAAT GAGACACCAG GGATTAGATA TCAGTACAAT

      2470       2480       2490       2500       2510       2520
GTGCTTCCAC AGGGATGGAA AGGATCACCA GCAATATTCC AAAGTAGCAT GACAAAAATC

      2530       2540       2550       2560       2570       2580
TTAGAGCCTT TTAGAAAACA AAATCCAGAC ATAGTTATCT ATCAATACAT GGATGATTTG

      2590       2600       2610       2620       2630       2640
TATGTAGGAT CTGACTTAGA AATAGGGCAG CATAGAACAA AAATAGAGGA GCTGAGACAA

      2650       2660       2670       2680       2690       2700
CATCTGTTGA GGTGGGGACT TACCACACCA GACAAAAAAC ATCAGAAAGA ACCTCCATTC

      2710       2720       2730       2740       2750       2760
CTTTGGATGG GTTATGAACT CCATCCTGAT AAATGGACAG TACAGCCTAT AGTGCTGCCA

      2770       2780       2790       2800       2810       2820
GAAAAAGACA GCTGGACTGT CAATGACATA CAGAAGTTAG TGGGAAAATT GAATTGGGCA

      2830       2840       2850       2860       2870       2880
AGTCAGATTT ACCCAGGGAT TAAAGTAAGG CAATTATGTA AACTCCTTAG AGGAACCAAA

      2890       2900       2910       2920       2930       2940
GCACTAACAG AAGTAATACC ACTAACAGAA GAAGCAGAGC TAGAACTGGC AGAAAACAGA

      2950       2960       2970       2980       2990       3000
GAGATTCTAA AAGAACCAGT ACATGGAGTG TATTATGACC CATCAAAAGA CTTAATAGCA

      3010       3020       3030       3040       3050       3060
GAAATACAGA AGCAGGGGCA AGGCCAATGG ACATATCAAA TTTATCAAGA GCCATTTAAA

      3070       3080       3090       3100       3110       3120
AATCTGAAAA CAGGAAAATA TGCAAGAACG AGGGGTGCCC ACACTAATGA TGTAAAACAA

      3130       3140       3150       3160       3170       3180
TTAACAGAGG CAGTGCAAAA AATAACCACA GAAAGCATAG TAATATGGGG AAAGACTCCT

      3190       3200       3210       3220       3230       3240
AAATTTAAAC TACCCATACA AAAGGAAACA TGGGAAACAT GGTGGACAGA GTATTGGCAA

      3250       3260       3270       3280       3290       3300
GCCACCTGGA TTCCTGAGTG GGAGTTTGTC AATACCCCTC CTTTAGTGAA ATTATGCTAC

      3310       3320       3330       3340       3350       3360
CAGTTAGAGA AAGAACCCAT AGTAGGAGCA GAAACGTTCT ATGTAGATGG GGCAGCTAGC

      3370       3380       3390       3400       3410       3420
AGGGAGACTA AATTAGGAAA AGCAGGATAT GTTACTAATA GAGGAAGACA AAAAGTTGTC

      3430       3440       3450       3460       3470       3480
ACCCTAACTG ACACAACAAA TCAGAAGACT GAGTTACAAG CAATTCATCT AGCTTTGCAG

      3490       3500       3510       3520       3530       3540
GATTCGGGAT TAGAAGTAAA TATAGTAACA GACTCACAAT ATGCATTAGG AATCATTCAA

      3550       3560       3570       3580       3590       3600
GCACAACCAG ATAAAAGTGA ATCAGAGTTA GTCAATCAAA TAATAGAGCA GTTAATAAAA

      3610       3620       3630       3640       3650       3660
```

111248

08 026736

Fig22

```
AAG[illegible]AAAA[illegible] TCTATCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG AAATGAACAA

      3670       3680       3690       3700       3710       3720
GTAGATAAAT TAGTCAGTGC TGGAATCAGG AAAGTACTAT TTTTAGATGG AATAGATAAG

      3730       3740       3750       3760       3770       3780
GCCCAAGATG AACATGAGAA ATATCACAGT AATTGGAGAG CAATGGCTAG TGATTTTAAC

      3790       3800       3810       3820       3830       3840
CTGCCACCTG TAGTAGCAAA AGAAATAGTA GCCAGCTGTG ATAAATGTCA GCTAAAAGGA

      3850       3860       3870       3880       3890       3900
GAAGCCATGC ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAACTAGA TTGTACACAT

      3910       3920       3930       3940       3950       3960
TTAGAAGGAA AAGTTATCCT GGTAGCAGTT CATGTAGCCA GTGGATATAT AGAAGCAGAA

      3970       3980       3990       4000       4010       4020
GTTATTCCAG CAGAAACAGG GCAGGAAACA GCATACTTTC TTTTAAAATT AGCAGGAAGA

      4030       4040       4050       4060       4070       4080
TGGCCAGTAA AAACAATACA TACAGACAAT GGCAGCAATT TCACCAGTAC TACGGTTAAG

      4090       4100       4110       4120       4130       4140
GCCGCCTGTT GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA TCCCCAAAGT

      4150       4160       4170       4180       4190       4200
CAAGGAGTAG TAGAATCTAT GAATAAAGAA TTAAAGAAAA TTATAGGCCA GGTAAGAGAT

      4210       4220       4230       4240       4250       4260
CAGGCTGAAC ATCTTAAGAC AGCAGTACAA ATGGCAGTAT TCATCCACAA TTTTAAAAGA

      4270       4280       4290       4300       4310       4320
AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG GAAAGAATAG TAGACATAAT AGCAACAGAC

      4330       4340       4350       4360       4370       4380
ATACAAACTA AAGAATTACA AAAACAAATT ACAAAAATTC AAAATTTTCG GGTTTATTAC

      4390       4400       4410       4420       4430       4440
AGGGACAGCA GAGATCCACT TTGGAAAGGA CCAGCAAAGC TCCTCTGGAA AGGTGAAGGG

      4450       4460       4470       4480       4490       4500
GCAGTAGTAA TACAAGATAA TAGTGACATA AAAGTAGTGC CAAGAAGAAA AGCAAAGATC

      4510       4520       4530       4540       4550       4560
ATTAGGGATT ATGGAAAACA GATGGCAGGT GATGATTGTG TGGCAAGTAG ACAGGATGAG

      4570       4580       4590       4600       4610       4620
GATTAGAACA TGGAAAAGTT TAGTAAAACA CCATATGTAT GTTTCAGGGA AAGCTAGGGG

      4630       4640       4650       4660       4670       4680
ATGGTTTTAT AGACATCACT ATGAAAGCCC TCATCCAAGA ATAAGTTCAG AAGTACACAT

      4690       4700       4710       4720       4730       4740
CCCACTAGGG GATGCTAGAT TGGTAATAAC AACATATTGG GGTCTGCATA CAGGAGAAAG

      4750       4760       4770       4780       4790       4800
AGACTGGCAT CTGGGTCAGG GAGTCTCCAT AGAATGGAGG AAAAAGAGAT ATAGCACACA

      4810       4820       4830       4840       4850       4860
AGTAGACCCT GAACTAGCAG ACCAACTAAT TCATCTGTAT TACTTTGACT GTTTTTCAGA

      4870       4880       4890       4900       4910       4920
```

44

Fig 23

```
CTCTGGTATA AGAAAGCCCT TATTAGGACA TATAGTTAGC CCTAGGTGTG AATATCAAGC

      4930       4940       4950       4960       4970       4980
AGGACATAAC AAGGTAGGAT CTCTACAATA CTTGGCACTA GCAGCATTAA TAACACCAAA

      4990       5000       5010       5020       5030       5040
AAAGATAAAG CCAGCTTTGC CTAGTGTTAC GAAACTGACA GAGGATAGAT GGAACAAGCC

      5050       5060       5070       5080       5090       5100
CCAGAAGACC AAGGGCCACA GAGGGAGCCA CACAATGAAT GGACACTAGA GCTTTTAGAG

      5110       5120       5130       5140       5150       5160
GAGCTTAAGA ATGAAGCTGT TAGACATTTT CCTAGGATTT GGCTCCATGG CTTAGGGCAA

      5170       5180       5190       5200       5210       5220
CATATCTATG AAACTTATGG GGATACTTGG GCAGGAGTGG AAGCCATAAT AAGAATTCTG

      5230       5240       5250       5260       5270       5280
CAACAACTGC TGTTTATCCA TTTCAGAATT GGGTGTCGAC ATAGCAGAAT AGGCGTTACT

      5290       5300       5310       5320       5330       5340
CAACAGAGGA GAGCAAGAAA TGGAGCCAGT AGATCCTAGA CTAGAGCCCT GGAAGCATCC

      5350       5360       5370       5380       5390       5400
AGGAAGTCAG CCTAAAACTG CTTGTACCAC TTGCTATTGT AAAAAGTGTT GCTTTCATTG

      5410       5420       5430       5440       5450       5460
CCAAGTTTGT TTCACAACAA AAGCCTTAGG CATCTCCTAT GGCAGGAAGA AGCGGAGACA

      5470       5480       5490       5500       5510       5520
GCGACGAAGA CCTCCTCAAG GCAGTCAGAC TCATCAAGTT TCTCTATCAA AGCAGTAAGT

      5530       5540       5550       5560       5570       5580
AGTACATGTA ATGCAACCTA TACAAATAGC AATAGCAGCA TTAGTAGTAG CAATAATAAT

      5590       5600       5610       5620       5630       5640
AGCAATAGTT GTGTGGTCCA TAGTAATCAT AGAATATAGG AAAATATTAA GACAAAGAAA

      5650       5660       5670       5680       5690       5700
AATAGACAGG TTAATTGATA GACTAATAGA AAGAGCAGAA GACAGTGGCA ATGAGAGTGA

      5710       5720       5730       5740       5750       5760
AGGAGAAATA TCAGCACTTG TGGAGATGGG GGTGGAAATG GGGCACCATG CTCCTTGGGA

      5770       5780       5790       5800       5810       5820
TATTGATGAT CTGTAGTGCT ACAGAAAAAT TGTGGGTCAC AGTCTATTAT GGGGTACCTG

      5830       5840       5850       5860       5870       5880
TGTGGAAGGA AGCAACCACC ACTCTATTTT GTGCATCAGA TGCTAAAGCA TATGATACAG

      5890       5900       5910       5920       5930       5940
AGGTACATAA TGTTTGGGCC ACACATGCCT GTGTACCCAC AGACCCCAAC CCACAAGAAG

      5950       5960       5970       5980       5990       6000
TAGTATTGGT AAATGTGACA GAAAATTTTA ACATGTGGAA AAATGACATG GTAGAACAGA

      6010       6020       6030       6040       6050       6060
TGCATGAGGA TATAATCAGT TTATGGGATC AAAGCCTAAA GCCATGTGTA AAATTAACCC

      6070       6080       6090       6100       6110       6120
CACTCTGTGT TAGTTTAAAG TGCACTGATT TGGGGAATGC TACTAATACC AATAGTAGTA

      6130       6140       6150       6160       6170       6180
```

08 026736

2824

```
ATACCAATAG TAGTAGCGGG GAAATGATGA TGGAGAAAGG AGAGATAAAA AACTGCTCTT
      6190       6200       6210       6220       6230       6240
TCAATATCAG CACAAGCATA AGAGGTAAGG TGCAGAAAGA ATATGCATTT TTTTATAAAC
      6250       6260       6270       6280       6290       6300
TTGATATAAT ACCAATAGAT AATGATACTA CCAGCTATAC GTTGACAAGT TGTAACACCT
      6310       6320       6330       6340       6350       6360
CAGTCATTAC ACAGGCCTGT CCAAAGGTAT CCTTTGAGCC AATTCCCATA CATTATTGTG
      6370       6380       6390       6400       6410       6420
CCCCGGCTGG TTTTGCGATT CTAAAATGTA ATAATAAGAC GTTCAATGGA ACAGGACCAT
      6430       6440       6450       6460       6470       6480
GTACAAATGT CAGCACAGTA CAATGTACAC ATGGAATTAG GCCAGTAGTA TCAACTCAAC
      6490       6500       6510       6520       6530       6540
TGCTGTTGAA TGGCAGTCTA GCAGAAGAAG AGGTAGTAAT TAGATCTGCC AATTTCACAG
      6550       6560       6570       6580       6590       6600
ACAATGCTAA AACCATAATA GTACAGCTGA ACCAATCTGT AGAAATTAAT TGTACAAGAC
      6610       6620       6630       6640       6650       6660
CCAACAACAA TACAAGAAAA AGTATCCGTA TCCAGAGGGG ACCAGGGAGA GCATTTGTTA
      6670       6680       6690       6700       6710       6720
CAATAGGAAA AATAGGAAAT ATGAGACAAG CACATTGTAA CATTAGTAGA GCAAAATGCA
      6730       6740       6750       6760       6770       6780
ATGCCACTTT AAAACAGATA GCTAGCAAAT TAAGAGAACA ATTTGGAAAT AATAAAACAA
      6790       6800       6810       6820       6830       6840
TAATCTTTAA GCAATCCTCA GGAGGGGACC CAGAAATTGT AACGCACAGT TTTAATTGTG
      6850       6860       6870       6880       6890       6900
GAGGGGAATT TTTCTACTGT AATTCAACAC AACTGTTTAA TAGTACTTGG TTTAATAGTA
      6910       6920       6930       6940       6950       6960
CTTGGAGTAC TGAAGGGTCA AATAACACTG AAGGAAGTGA CACAATCACA CTCCCATGCA
      6970       6980       6990       7000       7010       7020
GAATAAAACA ATTTATAAAC ATGTGGCAGG AAGTAGGAAA AGCAATGTAT GCCCCTCCCA
      7030       7040       7050       7060       7070       7080
TCAGCGGACA AATTAGATGT TCATCAAATA TTACAGGGCT GCTATTAACA AGAGATGGTG
      7090       7100       7110       7120       7130       7140
GTAATAACAA CAATGGGTCC GAGATCTTCA GACCTGGAGG AGGAGATATG AGGGACAATT
      7150       7160       7170       7180       7190       7200
GGAGAAGTGA ATTATATAAA TATAAAGTAG TAAAAATTGA ACCATTAGGA GTAGCACCCA
      7210       7220       7230       7240       7250       7260
CCAAGGCAAA GAGAAGAGTG GTGCAGAGAG AAAAAAGAGC AGTGGGAATA GGAGCTTTGT
      7270       7280       7290       7300       7310       7320
TCCTTGGGTT CTTGGGAGCA GCAGGAAGCA CTATGGGCGC ACGGTCAATG ACGCTGACGG
      7330       7340       7350       7360       7370       7380
TACAGGCCAG ACAATTATTG TCTGGTATAG TGCAGCAGCA GAACAATTTG CTGAGGGCTA
      7390       7400       7410       7420       7430       7440
```

46

08 026736

```
TTGAGGCGCA ACAGCATCTG TTGCAACTCA CAGTCTGGGG CATCAAGCAG CTCCAGGCAA

      7450       7460       7470       7480       7490       7500
GAATCCTGGC TGTGGAAAGA TACCTAAAGG ATCAACAGCT CCTGGGGATT TGGGGTTGCT

      7510       7520       7530       7540       7550       7560
CTGGAAAACT CATTTGCACC ACTGCTGTGC CTTGGAATGC TAGTTGGAGT AATAAATCTC

      7570       7580       7590       7600       7610       7620
TGGAACAGAT TTGGAATAAC ATGACCTGGA TGCAGTGGGA CAGAGAAATT AACAATTACA

      7630       7640       7650       7660       7670       7680
CAAGCTTAAT ACATTCCTTA ATTGAAGAAT CGCAAAACCA GCAAGAAAAG AATGAACAAG

      7690       7700       7710       7720       7730       7740
AATTATTGGA ATTAGATAAA TGGGCAAGTT TGTGGAATTG GTTTAACATA ACAAATTGGC

      7750       7760       7770       7780       7790       7800
TGTGGTATAT AAAAATATTC ATAATGATAG TAGGAGGCTT GGTAGGTTTA AGAATAGTTT

      7810       7820       7830       7840       7850       7860
TTGCTGTACT TTCTATAGTG AATAGAGTTA GGCAGGGATA TTCACCATTA TCGTTTCAGA

      7870       7880       7890       7900       7910       7920
CCCACCTCCC AACCCCGAGG GGACCCGACA GGCCCGAAGG AATAGAAGAA GAAGGTGGAG

      7930       7940       7950       7960       7970       7980
AGAGAGACAG AGACAGATCC ATTCGATTAG TGAACGGATC CTTAGCACTT ATCTGGGACG

      7990       8000       8010       8020       8030       8040
ATCTGCGGAG CCTTGTGCCT CTTCAGCTAC CACCGCTTGA GAGACTTACT CTTGATTGTA

      8050       8060       8070       8080       8090       8100
ACGAGGATTG TGGAACTTCT GGGACGCAGG GGGTGGGAAG CCCTCAAATA TTGGTGGAAT

      8110       8120       8130       8140       8150       8160
CTCCTACAGT ATTGGAGTCA GGAACTAAAG AATAGTGCTG TTAGCTTGCT CAATGCCACA

      8170       8180       8190       8200       8210       8220
GCCATAGCAG TAGCTGAGGG GACAGATAGG GTTATAGAAG TAGTACAAGG AGCTTGTAGA

      8230       8240       8250       8260       8270       8280
GCTATTCGCC ACATACCTAG AAGAATAAGA CAGGGCTTGG AAAGGATTTT GCTATAAGAT

      8290       8300       8310       8320       8330       8340
GGGTGGCAAG TGGTCAAAAA GTAGTGTGGT TGGATGGCCT ACTGTAAGGG AAAGAATGAG

      8350       8360       8370       8380       8390       8400
ACGAGCTGAG CCAGCAGCAG ATGGGGTGGG AGCAGCATCT CGAGACCTGG AAAAACATGG

      8410       8420       8430       8440       8450       8460
AGCAATCACA AGTAGCAATA CAGCAGCTAC CAATGCTGCT TGTGCCTGGC TAGAAGCACA

      8470       8480       8490       8500       8510       8520
AGAGGAGGAG GAGGTGGGTT TTCCAGTCAC ACCTCAGGTA CCTTTAAGAC CAATGACTTA

      8530       8540       8550       8560       8570       8580
CAAGGCAGCT GTAGATCTTA GCCACTTTTT AAAAGAAAAG GGGGGACTGG AAGGGCTAAT

      8590       8600       8610       8620       8630       8640
TCACTCCCAA CGAAGACAAG ATATCCTTGA TCTGTGGATC TACCACACAC AAGGCTACTT

      8650       8660       8670       8680       8690       8700
```

08 026736

```
CCCTGATTGG CAGAACTACA CACCAGGGCC AGGGGTCAGA TATCCACTGA CCTTTGGATG

      8710       8720       8730       8740       8750       8760
GTGCTACAAG CTAGTACCAG TTGAGCCAGA TAAGGTAGAA GAGGCCAATA AAGGAGAGAA

      8770       8780       8790       8800       8810       8820
CACCAGCTTG TTACACCCTG TGAGCCTGCA TGGAATGGAT GACCCTGAGA GAGAAGTGTT

      8830       8840       8850       8860       8870       8880
AGAGTGGAGG TTTGACAGCC GCCTAGCATT TCATCACGTG GCCCGAGAGC TGCATCCGGA

      8890       8900       8910       8920       8930       8940
GTACTTCAAG AACTGCTGAC ATCGAGCTTG CTACAAGGGA CTTTCCGCTG GGGACTTTCC

      8950       8960       8970       8980       8990       9000
AGGGAGGCGT GGCCTGGGCG GAACTGGGGA GTGGCGAGCC CTCAGATGCT GCATATAAGC

      9010       9020       9030       9040       9050       9060
AGCTGCTTTT TGCCTGTACT GGGTCTCTCT GGTTAGACCA GATTTGAGCC TGGGAGCTCT

      9070       9080       9090       9100          0          0
CTGGCTAACT AGGGAACCCA CTGCTTAAGC CTCAATAAAG CTT
```

Fig 26